



# STIC Search Report

## EIC 2100

**STIC Database Tracking Number: 112178**

**TO: Chongshan Chen**  
**Location:**  
**Art Unit : 2172**  
**Thursday, January 15, 2004**

**Case Serial Number: 09897803**

**From: Geoffrey St. Leger**  
**Location: EIC 2100**  
**PK2-4B30**  
**Phone: 308-7800**

**geoffrey.stleger@uspto.gov**

### Search Notes

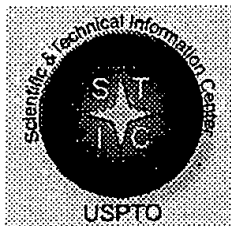
Dear Examiner Chen,

Attached please find the results of your search request for application 09897803. I searched Dialog, ACM, IEEE and the Internet.

Please let me know if you have any questions.

Regards,

Geoffrey St. Leger  
4B30/308-7800



# STIC EIC 2100 112178

## Search Request Form

Today's Date:

1-15-04

What date would you like to use to limit the search?

Priority Date: 7/2/01

Other:

Name Chongshan Chen

AU 2172 Examiner # 79547

Room # 4B25 Phone 305-8319

Serial # 09/897,803

Format for Search Results (Circle One):

☒ PAPER ☐ DISK ☐ EMAIL

Where have you searched so far?

☒ USP ☐ DWPI ☐ EPO ☐ JPO ☐ ACM ☐ IBM TDB

IEEE ☐ INSPEC ☐ SPI ☐ Other \_\_\_\_\_

Is this a "Fast & Focused" Search Request? (Circle One) ☒ YES ☐ NO

A "Fast & Focused" Search is completed in 2-3 hours (maximum). The search must be on a very specific topic and meet certain criteria. The criteria are posted in EIC2100 and on the EIC2100 NPL Web Page at <http://ptoweb/patents/stic/stic-tc2100.htm>.

What is the topic, novelty, motivation, utility, or other specific details defining the desired focus of this search? Please include the concepts, synonyms, keywords, acronyms, definitions, strategies, and anything else that helps to describe the topic. Please attach a copy of the abstract, background, brief summary, pertinent claims and any citations of relevant art you have found.

A method for database administration and replication, comprising the steps of:

Proving a database management system with an integrated (built in) random sampling facility;

Selecting a default sample size value S;

Selectively receiving a desired sample size value D and setting said default sample size value S to said desired sample size value D when said desired sample size value D is received;

Randomly sampling S records of the database using said random sampling facility;

Wherein the step of sampling said S records includes randomly sampling the S records utilizing dataspaces including:

At least one index dataspace;

At least one key dataspace; and,

At least one statistics dataspace;

Storing statistics for each of said S records, wherein said statistics include a record key for each record; and

Producing at least one of:

An extrapolated replication partition analysis based on said statistics; and

A partial replication partition analysis based on said statistics.

Wherein the step of producing at least one of said partition analysis includes the step of defining multiple partition boundaries;

STIC Searcher Geoffrey St. Leger Phone 308-7800

Date picked up 1/15/4 Date Completed 1/15/4



File 411:DIALINDEX(R)

DIALINDEX(R)

(c) 2004 The Dialog Corporation plc

\*\*\* DIALINDEX search results display in an abbreviated \*\*\*

\*\*\* format unless you enter the SET DETAIL ON command. \*\*\*

?set file sall

>>>"SALL" is not a valid Dialindex category

>>>No valid files specified

?set files all

You have 556 files in your file list.

(To see banners, use SHOW FILES command)

?s (random??(3n)sampl???) (100n) (dataspace? ? or data()space? ?)

Your SELECT statement is:

s (random??(3n)sampl???) (100n) (dataspace? ? or data()space? ?)

Items	File
----	----
3	2: INSPEC_1969-2004/Jan W1
1	7: Social_SciSearch(R)_1972-2004/Jan W2
1	8: Ei Compendex(R)_1970-2004/Jan W1
3	34: SciSearch(R) Cited Ref Sci_1990-2004/Jan W2
Examined 50 files	
1	94: JICST-EPlus_1985-2004/Jan W1
1	144: Pascal_1973-2004/Jan W1
Examined 100 files	
1	148: Gale Group Trade & Industry DB_1976-2004/Jan 15
1	155: MEDLINE(R)_1966-2004/Jan W2
1	180: Federal Register_1985-2004/Jan 14
Examined 150 files	
Examined 200 files	
2	340: CLAIMS(R)/US Patent_1950-03/Jan 13
1	348: EUROPEAN PATENTS_1978-2004/Jan W02
3	349: PCT FULLTEXT_1979-2002/UB=20031225,UT=20031218
Examined 250 files	
3	440: Current Contents Search(R)_1990-2004/Jan 15
Examined 300 files	
Examined 350 files	
Examined 400 files	
Processing	
Processing	
6	654: US Pat.Fulll._1976-2004/Jan 13
Examined 450 files	
Examined 500 files	
Examined 550 files	

14 files have one or more items; file list includes 556 files.

File 2:INSPEC 1969-2004/Jan W1  
(c) 2004 Institution of Electrical Engineers  
File 7:Social SciSearch(R) 1972-2004/Jan W2  
(c) 2004 Inst for Sci Info  
File 8:Ei Compendex(R) 1970-2004/Jan W1  
(c) 2004 Elsevier Eng. Info. Inc.  
File 34:SciSearch(R) Cited Ref Sci 1990-2004/Jan W2  
(c) 2004 Inst for Sci Info  
File 94:JICST-EPlus 1985-2004/Jan W1  
(c)2004 Japan Science and Tech Corp(JST)  
File 144:Pascal 1973-2004/Jan W1  
(c) 2004 INIST/CNRS  
File 148:Gale Group Trade & Industry DB 1976-2004/Jan 15  
(c)2004 The Gale Group  
File 155:MEDLINE(R) 1966-2004/Jan W2  
(c) format only 2004 The Dialog Corp.  
File 180:Federal Register 1985-2004/Jan 14  
(c) 2004 format only The DIALOG Corp  
File 340:CLAIMS(R)/US Patent 1950-03/Jan 13  
(c) 2004 IFI/CLAIMS(R)  
File 348:EUROPEAN PATENTS 1978-2004/Jan W02  
(c) 2004 European Patent Office  
File 349:PCT FULLTEXT 1979-2002/UB=20031225,UT=20031218  
(c) 2003 WIPO/Univentio  
File 440:Current Contents Search(R) 1990-2004/Jan 15  
(c) 2004 Inst for Sci Info  
File 654:US Pat.Full. 1976-2004/Jan 13  
(c) Format only 2004 The Dialog Corp.

Set	Items	Description
S1	28	(RANDOM??(3N) SAMPL???) (100N) (DATASPACE? ? OR DATA()SPACE? - ?)
S2	18	RD (unique items)

2/5/3 (Item 3 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2004 Institution of Electrical Engineers. All rts. reserv.

5034359 INSPEC Abstract Number: B9510-6140C-469, C9510-1250-239

**Title: Registration of 3-D images by genetic optimization**

Author(s): Jacq, J.-J.; Roux, C.

Author Affiliation: Dept. Image et Traitment de l'Inf., Telecom Bretagne, Brest, France

Journal: Pattern Recognition Letters vol.16, no.8 p.823-41

Publication Date: Aug. 1995 Country of Publication: Netherlands

CODEN: PRLEDG ISSN: 0167-8655

U.S. Copyright Clearance Center Code: 0167-8655/95/\$09.50

Language: English Document Type: Journal Paper (JP)

Treatment: Applications (A); Theoretical (T)

**Abstract:** We present a framework for solving the 3D registration problem in medical imaging based on a canonical genetic algorithm (CGA). The issue of 3D registration is stated as an optimization problem in both application cases presented, i.e., volume-to-volume and surface-to-volume registration. The CGA uses a stochastic fitness function which operates on **randomly selected samples** of the **data space**. At a higher level, an adaptive search space scaling technique is presented which operates by successive activations of the CGA procedure. The former features ensure a lower complexity of the search algorithm and a good accuracy of the final solution. Volume-to-volume and surface-to-volume registration are then considered. The features that are specific to the application (the actual optimization space, the fitness or distance function, the GA parameters) are introduced. Results concerning two registration problems using 3D computerized tomography data are presented and discussed. (17 Refs)

Subfile: B C

Descriptors: biomedical imaging; computerised tomography; genetic algorithms; image reconstruction; image registration; medical image processing; search problems; stereo image processing; tomography

Identifiers: 3D image registration; medical imaging; canonical genetic algorithm; 3D computerized tomograph; optimization; volume-to-volume registration; surface-to-volume registration; stochastic fitness function; data space; adaptive search space scaling; search algorithm; shape reconstruction

Class Codes: B6140C (Optical information, image and video signal processing); B7510 (Biomedical measurement and imaging); B0260 (Optimisation techniques); C1250 (Pattern recognition); C7330 (Biology and medical computing); C5260B (Computer vision and image processing techniques); C1180 (Optimisation techniques)

Copyright 1995, IEE

2/5/4 (Item 1 from file: 94)  
DIALOG(R)File 94:JICST-EPlus  
(c)2004 Japan Science and Tech Corp(JST). All rts. reserv.

04727330 JICST ACCESSION NUMBER: 01A0216796 FILE SEGMENT: PreJICST-E  
**Data Sampling for Evaluation of Structural Diversity of Chemical Compounds.**  
TAKEZAWA HIROSHI (1); TAKAHASHI YOSHIMASA (1)  
(1) Toyohashi Univ. of Technol.  
Joho Kagaku Toronkai, Kozo Kassei Sokan Shinpojiumu Koen Yoshishu, 2000,  
VOL.23rd-28th, PAGE.208-211  
JOURNAL NUMBER: X0081AAK  
LANGUAGE: Japanese COUNTRY OF PUBLICATION: Japan  
DOCUMENT TYPE: Conference Proceeding  
MEDIA TYPE: Printed Publication

ABSTRACT: This paper describes a data sampling method for the evaluation of structural diversity of chemical compounds. Three different types of methods ( **random sampling** , cell partitioning method and clustering-based method) were investigated using a trial set of 5000 points prepared by two-dimensional random numbers. For cell partitioning method, two different approaches were tested: the sample distribution density was taken account for one, and not for the other. The results showed that the cell partitioning method with taking the density gives the most diverse sampling on that space. The method was applied to diverse sampling of chemical structures on a higher dimensional structural feature space characterized by topological fragment spectra. For this case, data sampling was carried out on a reduced **data space** that is produced by mathematical mapping. The result also validated the usability of the cell partitioning approach combined with the space reduction. (author abst.)

2/3,K/1 (Item 1 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2004 Institution of Electrical Engineers. All rts. reserv.

6908985 INSPEC Abstract Number: C2001-06-6130-001

**Title: Nonlinear mapping of massive data sets by fuzzy clustering and neural networks**

Author(s): Rassokhin, D.N.; Lobanov, V.S.; Agrafiotis, D.K.  
Author Affiliation: 3-Dimensional Pharm. Inc., Exton, PA, USA  
Journal: Journal of Computational Chemistry vol.22, no.4 p.373-86  
Publisher: Wiley,  
Publication Date: March 2001 Country of Publication: USA  
CODEN: JCCHDD ISSN: 0192-8651  
SICI: 0192-8651(200103)22:4L:373:NMMD;1-0  
Material Identity Number: J333-2001-003  
Language: English  
Subfile: C  
Copyright 2001, IEE

...Abstract: to relatively small data sets. We recently demonstrated that nonlinear maps derived from a small **random sample** of a large data set exhibit the same structure and characteristics as that of the...

...algorithm based on local learning. The method employs a fuzzy clustering methodology to partition the **data space** into a set of Voronoi polyhedra, and uses a separate neural network to perform the...

2/3,K/2 (Item 2 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2004 Institution of Electrical Engineers. All rts. reserv.

6829357 INSPEC Abstract Number: C2001-03-7320-019

**Title: Nonlinear mapping networks**

Author(s): Agrafiotis, D.K.; Lobanov, V.S.  
Author Affiliation: 3-Dimensional Pharm. Inc., Exton, PA, USA  
Journal: Journal of Chemical Information and Computer Sciences vol.40, no.6 p.1356-62  
Publisher: ACS,  
Publication Date: Nov.-Dec. 2000 Country of Publication: USA  
CODEN: JCISD8 ISSN: 0095-2338  
SICI: 0095-2338(200011/12)40:6L:1356:NMN;1-9  
Material Identity Number: J263-2000-006  
U.S. Copyright Clearance Center Code: 0095-2338/2000/\$19.00  
Language: English  
Subfile: C  
Copyright 2001, IEE

...Abstract: unique for their conceptual simplicity and ability to reproduce the topology and structure of the **data space** in a faithful and unbiased manner. However, a major shortcoming of these methods is their

... the principle of probability sampling, the method employs a classical algorithm to project a small **random sample**, and then "learns" the underlying nonlinear transform using a multilayer neural network trained with the...

2/3,K/3 (Item 3 from file: 2)  
DIALOG(R)File 2:INSPEC  
(c) 2004 Institution of Electrical Engineers. All rts. reserv.

5034359 INSPEC Abstract Number: B9510-6140C-469, C9510-1250-239

**Title: Registration of 3-D images by genetic optimization**

Author(s): Jacq, J.-J.; Roux, C.  
Author Affiliation: Dept. Image et Traitement de l'Inf., Telecom Bretagne, Brest, France  
Journal: Pattern Recognition Letters vol.16, no.8 p.823-41

Publication Date: Aug. 1995 Country of Publication: Netherlands  
CODEN: PRLEDG ISSN: 0167-8655  
U.S. Copyright Clearance Center Code: 0167-8655/95/\$09.50  
Language: English  
Subfile: B C  
Copyright 1995, IEE

...Abstract: and surface-to-volume registration. The CGA uses a stochastic fitness function which operates on **randomly** selected **samples** of the **data space**. At a higher level, an adaptive search space scaling technique is presented which operates by...

2/3,K/4 (Item 1 from file: 94)  
DIALOG(R)File 94:JICST-Eplus  
(c)2004 Japan Science and Tech Corp(JST). All rts. reserv.

04727330 JICST ACCESSION NUMBER: 01A0216796 FILE SEGMENT: PreJICST-E  
**Data Sampling for Evaluation of Structural Diversity of Chemical Compounds.**  
TAKEZAWA HIROSHI (1); TAKAHASHI YOSHIMASA (1)  
(1) Toyohashi Univ. of Technol.  
Joho Kagaku Toronkai, Kozo Kassei Soka Shinpojiumu Koen Yoshishu, 2000,  
VOL.23rd-28th, PAGE.208-211  
JOURNAL NUMBER: X0081AAK  
LANGUAGE: Japanese COUNTRY OF PUBLICATION: Japan  
DOCUMENT TYPE: Conference Proceeding  
MEDIA TYPE: Printed Publication

...ABSTRACT: method for the evaluation of structural diversity of chemical compounds. Three different types of methods ( **random sampling** , cell partitioning method and clustering-based method) were investigated using a trial set of 5000...

...by topological fragment spectra. For this case, data sampling was carried out on a reduced **data space** that is produced by mathematical mapping. The result also validated the usability of the cell...

2/3,K/5 (Item 1 from file: 148)  
DIALOG(R)File 148:Gale Group Trade & Industry DB  
(c)2004 The Gale Group. All rts. reserv.

16071212 SUPPLIER NUMBER: 101941076 (USE FORMAT 7 OR 9 FOR FULL TEXT )  
**Current labor statistics.**  
Monthly Labor Review, 126, 3, 31(66)  
March, 2003  
ISSN: 0098-1818 LANGUAGE: English RECORD TYPE: Fulltext  
WORD COUNT: 28018 LINE COUNT: 08590

... Injuries and Illnesses  
Description of the series  
The Survey of Occupational Injuries and Illnesses collects **data** from employers about their workers' job-related nonfatal injuries and illnesses. The information that employers...

...Federal-State cooperative program with an independent sample selected for each participating State. A stratified **random sample** with a Neyman allocation is selected to represent all private industries in the State. The...

2/3,K/6 (Item 1 from file: 180)  
DIALOG(R)File 180:Federal Register  
(c) 2004 format only The DIALOG Corp. All rts. reserv.

DIALOG Accession Number: 02274118 Supplier Number: 930201997



2/3,K/7 (Item 1 from file: 340)  
DIALOG(R)File 340:CLAIMS(R)/US Patent  
(c) 2004 IFI/CLAIMS(R). All rts. reserv.

10260572 2003-0004973

**E/RANDOM SAMPLING AS A BUILT-IN FUNCTION FOR DATABASE ADMINISTRATION AND REPLICATION**

Inventors: Harper John William (US); Slishman Gordon Robert (US)  
Assignee: International Business Machines Corp  
Assignee Code: 42640

	Kind	Publication Number	Date	Application Number	Date
Priority Applic:	A1	US 20030004973	20030102	US 2001897803	20010702
				US 2001897803	20010702

Non-exemplary Claims: ...as set forth in claim 6, wherein the step of sampling said S records includes **randomly sampling** the S records utilizing **dataspaces** including: at least one index **dataspace** ; at least one key **dataspace** ; and, at least one statistics **dataspace** .  
...

...15. A database management system (DBMS) for managing an associated database, the DBMS comprising: **random sampling** facility integrated with the database management system; first database analysis tools using said integrated **random sampling** facility for generating extrapolated reports on database content; second database analysis tools using said integrated **random sampling** facility for generating extrapolated reports on database size; and, database replication tools adapted to execute

2/3,K/8 (Item 2 from file: 340)  
DIALOG(R)File 340:CLAIMS(R)/US Patent  
(c) 2004 IFI/CLAIMS(R). All rts. reserv.

10260543 2003-0004944

**E/PARTITION BOUNDARY DETERMINATION USING RANDOM SAMPLING ON VERY LARGE DATABASES**

Inventors: Harper John William (US); Slishman Gordon Robert (US)  
Assignee: International Business Machines Corp  
Assignee Code: 42640

	Kind	Publication Number	Date	Application Number	Date
Priority Applic:	A1	US 20030004944	20030102	US 2001897853	20010702
				US 2001897853	20010702

Non-exemplary Claims: ...12. The method as set forth in claim 1, wherein the step of **randomly sampling** said S records includes **randomly sampling** the S records utilizing **dataspaces** including: at least one index **dataspace** ; at least one key **dataspace** ; and, at least one statistics **dataspace** .  
...

...program routine having a random number generating algorithm; a second computer program routine having a **random sampling** facility utilizing said first program routine to randomly read records from a database and store

2/3,K/9 (Item 1 from file: 348)  
DIALOG(R)File 348:EUROPEAN PATENTS  
(c) 2004 European Patent Office. All rts. reserv.

01067933

**WEIGHTLESS BINARY N-TUPLE THRESHOLDING HIERARCHIES**  
**HIERARCHISCHE STRUKTUR ZUR SCHWELLENVERGLEICHUNG UNGEWOGENER BINARDATEN**  
**HIERARCHIES DE DEFINITION DE SEUILS N-TUPLES, BINAIRE ET SANS POIDS**  
PATENT ASSIGNEE:

BAE SYSTEMS plc, (427897), Warwick House, P.O. Box 87, Farnborough  
Aerospace Centre, Farnborough, Hampshire GU14 6YU, (GB), (Proprietor  
designated states: all)

INVENTOR:

KING, Douglas Beverley Stevenson, British A.M.A.andA, Electr.Eng.,W354B,  
W.Aerodrome, Warton,Nr Preston, Lancs. PR4 1AX, (GB)  
MACDIARMID, Ian Peter, British A.M.A.and Aero, Elect.,W423, W. Aerodrome,  
Warton,Nr Preston, Lancs. PR4 1AX, (GB)  
MOORE, Colin, British A.M.A.and A, elec.Eng.,W354B , W.Aerodrome,  
Warton,Nr Preston, Lancs. PR4 1AX, (GB)

LEGAL REPRESENTATIVE:

Newell, William Joseph (53194), Wynne-Jones, Laine & James 22 Rodney Road  
, Cheltenham Gloucestershire GL50 1JJ, (GB)

PATENT (CC, No, Kind, Date): EP 1040408 A1 001004 (Basic)  
EP 1040408 B1 020821  
WO 99032962 990701

APPLICATION (CC, No, Date): EP 98962564 981218; WO 98GB3837 981218

PRIORITY (CC, No, Date): GB 9726752 971219; GB 9823382 981027

DESIGNATED STATES: DE; ES; FR; GB; IT; NL; SE

INTERNATIONAL PATENT CLASS: G06F-007/02; G06F-015/80

NOTE:

No A-document published by EPO

LANGUAGE (Publication,Procedural,Application): English; English; English  
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	200234	607
CLAIMS B	(German)	200234	557
CLAIMS B	(French)	200234	695
SPEC B	(English)	200234	2376
Total word count - document A			0
Total word count - document B			4235
Total word count - documents A + B			4235

...SPECIFICATION is practically viable.

It is important to randomly connect the pattern matcher outputs between  
the **data space** and the hierarchical structure because the pattern  
matchers can often "clump" results, e.g. 1111111011000000001...

...of the first layer sum and threshold devices 14 of Figure 2 is seen to  
**randomly sample the data space**.

There may be certain applications that do not want this random mapping  
- in general, if...

2/3,K/10 (Item 1 from file: 349)  
DIALOG(R)File 349:PCT FULLTEXT  
(c) 2003 WIPO/Univentio. All rts. reserv.

00837972 \*\*Image available\*\*

**SYSTEM, METHOD, AND COMPUTER PROGRAM PRODUCT FOR REPRESENTING OBJECT**  
**RELATIONSHIPS IN A MULTIDIMENSIONAL SPACE**  
**SYSTEME, PROCEDE ET PROGICIEL POUR LA REPRESENTATION DE RELATIONS ENTRE**  
**OBJETS DANS UN ESPACE MULTIDIMENSIONNEL**

Patent Applicant/Assignee:

3-DIMENSIONAL PHARMACEUTICALS INC, Eagleview Corporate Center, Suite 104,  
665 Stockton Drive, Exton, PA 19341, US, US (Residence), US  
(Nationality), (For all designated states except: US)

Patent Applicant/Inventor:

AGRAFIOTIS Dimitris K, 660 Perimeter Drive, Downingtown, PA 19335, US, US  
(Residence), US (Nationality), (Designated only for: US)  
RASSOKHIN Dmitrii N, 101 Parker Court, Exton, PA 19341, US, US  
(Residence), RU (Nationality), (Designated only for: US)  
LOBANOV Victor S, 815 Azalea Drive, North Brunswick, NJ 08902, US, US  
(Residence), RU (Nationality), (Designated only for: US)  
SALEMME F Raymond, 1970 Timber Lakes Drive, Yardley, PA 19067, US, US  
(Residence), US (Nationality), (Designated only for: US)

Legal Representative:

LEE Michael Q (et al) (agent), Sterne, Kessler, Goldstein, & Fox  
P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC  
20005-3934, US,

Patent and Priority Information (Country, Number, Date):

Patent: WO 200171624 A1 20010927 (WO 0171624)  
Application: WO 2001US8974 20010322 (PCT/WO US0108974)  
Priority Application: US 2000191108 20000322

Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU  
CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR  
KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE  
SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW  
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR  
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG  
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW  
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 10642

Fulltext Availability:

Detailed Description

Detailed Description

... The approach employs an iterative algorithm based on subset  
refinements to nonlinearly map a small  
**random sample** which reflects the overall structure of the data, and  
then  
"learns" the underlying nonlinear transform...

...networks, each specializing in a particular domain of the feature space.  
The  
partitioning of the **data space** can be carried out using a

2/3,K/11 (Item 2 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2003 WIPO/Univentio. All rts. reserv.

00753772 \*\*Image available\*\*

METHOD, SYSTEM AND COMPUTER PROGRAM PRODUCT FOR NON-LINEAR MAPPING OF  
MULTI-DIMENSIONAL DATA  
PROCEDE, SYSTEME ET PROGRAMME INFORMATIQUE D'APPLICATION NON LINEAIRE DE  
DONNEES MULTIDIMENSIONNELLES

Patent Applicant/Assignee:

3-DIMENSIONAL PHARMACEUTICALS INC, Eagleview Corporate Center, Suite 104,  
665 Stockton Drive, Exton, PA 19341, US, US (Residence), US  
(Nationality)

Inventor(s):

AGRAFIOTIS Dimitris K, 660 Perimeter Drive, Downingtown, PA 19335, US  
LOBANOV Victor S, 24305 Cornerstone Drive, Yardley, PA 19067, US  
SALEMME Francis R, 1970 Timber Lakes, Yardley, PA 19067, US

Legal Representative:

LEE Michael Q, Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100  
New York Avenue, N.W., Washington, DC 20005-3934, US

Patent and Priority Information (Country, Number, Date):

Patent: WO 200067148 A1 20001109 (WO 0067148)  
Application: WO 2000US11838 20000503 (PCT/WO US0011838)  
Priority Application: US 99303671 19990503

Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY CA CH CN CR CU CZ DE

DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC  
LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK  
SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW  
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE  
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG  
(AP) GH GM KE LS MW SD SL SZ TZ UG ZW  
(EA) AM AZ BY KG KZ MD RU TJ TM

Publication Language: English

Filing Language: English

Fulltext Word Count: 14837

Fulltext Availability:

Detailed Description

Detailed Description

... alone for their conceptual elegance and ability to reproduce the  
topology and structure of the **data space** in a faithful and unbiased  
manner. Unfortunately, all  
known algorithms exhibit quadratic time complexity which...

...principle of

probability sampling, the method employs an algorithm to  
multi-dimensionally

scale a small **random sample**, and then "learns" the underlying  
non-linear

1 5 transform using a multi-layer perceptron...

2/3,K/12 (Item 3 from file: 349)

DIALOG(R)File 349:PCT FULLTEXT

(c) 2003 WIPO/Univentio. All rts. reserv.

00501610 \*\*Image available\*\*

**WEIGHTLESS BINARY N-TUPLE THRESHOLDING HIERARCHIES**

**HIERARCHIES DE DEFINITION DE SEUILS N-TUPLES, BINAIRES ET SANS POIDS**

Patent Applicant/Assignee:

BRITISH AEROSPACE PUBLIC LIMITED COMPANY,  
KING Douglas Beverley Stevenson;;  
MACDIARMID Ian Peter;;  
MOORE Colin;;

Inventor(s):

KING Douglas Beverley Stevenson;;  
MACDIARMID Ian Peter;;  
MOORE Colin;;

Patent and Priority Information (Country, Number, Date):

Patent: WO 9932962 A1 19990701

Application: WO 98GB3837 19981218 (PCT/WO GB9803837)

Priority Application: GB 9726752 19971219; GB 9823382 19981027

Designated States: AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES

FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU

LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA

UG US UZ VN YU ZW GH GM KE LS MW SD SZ UG ZW AM AZ BY KG KZ MD RU TJ TM

AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE BF BJ CF CG CI CM

GA GN GW ML MR NE SN TD TG

Publication Language: English

Fulltext Word Count: 3327

Fulltext Availability:

Detailed Description

Detailed Description

... is practically viable.

It is important to randomly connect the pattern matcher  
outputs between the **data space** and the hierarchical struc  
ture because the pattern matchers can often "clump" results,  
e.g...of the first layer sum and threshold devices 14  
of Figure 2 is seen to **randomly sample** the **data space**.

There may be certain applications that do not want this random mapping - in general, if...

2/3,K/13 (Item 1 from file: 654)

DIALOG(R)File 654:US Pat.Full.

(c) Format only 2004 The Dialog Corp. All rts. reserv.

0005401453 \*\*IMAGE Available

Derwent Accession: 1998-287200

**Method, system and computer program product for non-linear mapping of multi-dimensional data**

Inventor: Agrafiotis, Dimitris, INV

Lobanov, Victor, INV

Salemme, Francis, INV

Correspondence Address: STERNE, KESSLER, GOLDSTEIN & FOX PLLC, 1100 NEW YORK AVENUE, N.W., WASHINGTON, DC, 20005, US

	Publication Number	Kind	Date	Application Number	Filing Date
Main Patent	US 20030195897	A1	20031016	US 2003428075	20030502
Continuation	US 6571227			US 99303671	19990503
CIP	US 6453246			US 9873845	19980507
CIP	US 6295514			US 97963872	19971104
Provisional				US 60-30187	19961104

Fulltext Word Count: 14856

Description of the Invention:

...alone for their conceptual elegance and ability to reproduce the topology and structure of the **data space** in a faithful and unbiased manner. Unfortunately, all known algorithms exhibit quadratic time complexity which...

...principle of probability sampling, the method employs an algorithm to multi-dimensionally scale a small **random sample**, and then "learns" the underlying non-linear transform using a multi-layer perceptron trained with...

2/3,K/14 (Item 2 from file: 654)

DIALOG(R)File 654:US Pat.Full.

(c) Format only 2004 The Dialog Corp. All rts. reserv.

0005152651 \*\*IMAGE Available

Derwent Accession: 2003-352903

**Random sampling as a built-in function for database administration and replication**

Inventor: John Harper, INV

Gordon Slishman, INV

Assignee: International Business Machines Corporation (02)

Correspondence Address: Michael E. Hudzinski FAY, SHARPE, FAGAN,, MINNICH & McKEE, LLP 1100 Superior Avenue, Seventh Floor, Cleveland, OH, 44110-2518, US

	Publication Number	Kind	Date	Application Number	Filing Date
Main Patent	US 20030004973	A1	20030102	US 2001897803	20010702

Fulltext Word Count: 7426

Summary of the Invention:

...0028] Raw partition analysis, without **random sampling** analysis, places a heavy strain on a computer system in terms of memory usage and

typically requires multiple **dataspaces** . Random **sampling** relieves the strain on the computer system in terms of processing and memory requirements. Much less memory is required to analyze 20,000 **sampled** records using the **random sampling** approach than to analyze 2,000,000,000 records without sampling. However, in order to...

...with an unsampled approach which may be desirable under some circumstances, the preferred method using **random sampling** analysis utilizes one or more of each of the following types of **dataspaces** : index, key and statistics...

Description of the Invention:

...contain up to 8 gigabytes (GB) in keys, on a computer system having RAM 20 **dataspaces** of up to 2<sup>31</sup> bytes (2 GB), four **dataspaces** are required to store the keys. Another 2 GB are sufficient to store indices to the keys. However, the record statistics, even when compressed, may require dozens of **dataspaces** . To minimize the effort of storing and sorting, the present invention **randomly samples** a database and produces an extrapolated partition analysis 24 providing sufficiently accurate results. Preferably, the sample size selected is sufficiently small so that three **dataspaces** will suffice, one each for indices, keys, and statistics...

...0044] An analysis program 16, in communication with DBMS 14, partitions a **random sample** size of S records, and then scales the tabulated numbers by the ratio of the...0059] The memory required by a partition analysis, even when **random sampling** is employed, can be large and, consequently, multiple **dataspaces** may be required. For databases organized with indexes and keys, sampling may require one or more **dataspaces** , e.g. one or more index **dataspaces** , one or more key **dataspaces** , and one or more statistics **dataspaces** .  
[...]

...0060] After the **random sampling** has been performed by sampling facility 26, and analysis program 18 has performed a partition

Non-exemplary or Dependent Claim(s):

...as set forth in claim 6, wherein the step of sampling said S records includes **randomly sampling** the S records utilizing **dataspaces** including: at least one index **dataspace** ; at least one key **dataspace** ; and, at least one statistics **dataspace** .  
...

...sample size and setting said number S equal to said particular number; a means for **randomly sampling** S records of the database using said random sampling facility; a means for storing statistics

2/3,K/15 (Item 3 from file: 654)  
DIALOG(R)File 654:US Pat.Full.  
(c) Format only 2004 The Dialog Corp. All rts. reserv.

0005152622 \*\*IMAGE Available  
Derwent Accession: 2003-352889  
**Partition boundary determination using random sampling on very large databases**  
Inventor: John Harper, INV  
Gordon Sliselman, INV  
Assignee: International Business Machines Corporation (02)  
Correspondence Address: Michael E. Hudzinski FAY, SHARPE, FAGAN, MINNICH & McKEE, LLP, Seventh Floor 1100 Superior Avenue, Cleveland, OH, 44110-2518, US

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 20030004944	A1	20030102	US 2001897853	20010702

Fulltext Word Count: 7785

Summary of the Invention:

0027] Raw partition analysis, without **random sampling** analysis, places a heavy strain on a computer system in terms of memory usage and typically requires multiple **dataspaces**. **Random sampling** relieves the strain on the computer system in terms of processing and memory requirements. Much less memory is required to analyze 20,000 **sampled** records using the **random sampling** approach than to analyze 2,000,000,000 records without sampling. However, in order to...

...with an unsampled approach which may be desirable under some circumstances, the preferred method using **random sampling** analysis utilizes one or more of each of the following types of **dataspaces**: index, key and statistics...

Description of the Invention:

...contain up to 8 gigabytes (GB) in keys, on a computer system having RAM 18 **dataspaces** of up to 2<sup>31</sup> bytes (2 GB), four **dataspaces** are required to store the keys. Another 2 GB are sufficient to store indices to the keys. However, the record statistics, even when compressed, may require dozens of **dataspaces**. To minimize the effort of storing and sorting, the present invention **randomly samples** a database and produces an extrapolated partition analysis 22 providing sufficiently accurate results. Preferably, the sample size selected is sufficiently small so that three **dataspaces** will suffice, one each for indices, keys, and statistics...

...0044] An analysis program 16, in communication with the DBMS 14, partitions a **random sample** size of S records, and then scales the tabulated numbers by the ratio of the...0080] It should be realized that the memory required by a partition analysis, even when **random sampling** is employed can be large and, consequently, multiple **dataspaces** may be required. For databases organized with indexes and keys, sampling may require one or more **dataspaces**, e.g. one or more index **dataspaces**, one or more key **dataspaces**, and one or more statistics **dataspaces**. [...]

...0081] After **random sampling** has been performed by either sampling method, and analysis program 16 has performed necessary partition

Non-exemplary or Dependent Claim(s):

...12. The method as set forth in claim 1, wherein the step of **randomly sampling** said S records includes **randomly sampling** the S records utilizing **dataspaces** including: at least one index **dataspace**; at least one key **dataspace**; and, at least one statistics **dataspace**

2/3,K/16 (Item 4 from file: 654)  
DIALOG(R)File 654:US Pat.Full.  
(c) Format only 2004 The Dialog Corp. All rts. reserv.

0005036962 \*\*IMAGE Available  
Derwent Accession: 2002-025764

**System, method, and computer program product for representing object relationships in a multidimensional space**

Inventor: Dimitris Agrafiotis, INV  
Dmitrii Rassokhin, INV  
Victor Lobanov, INV  
F. Salemme, INV

Correspondence Address: STERNE, KESSLER, GOLDSTEIN & FOX PLLC, 1100 NEW YORK AVENUE, N.W., SUITE 600, WASHINGTON, DC, 20005-3934, US

	Number	Kind	Date	Number	Date
	-----	--	-----	-----	-----
Main Patent	US 20020091655	A1	20020711	US 2001814160	20010322
Provisional				US 60-191108	20000322

Fulltext Word Count: 12439

#### Description of the Invention:

...The approach employs an iterative algorithm based on subset refinements to nonlinearly map a small **random sample** which reflects the overall structure of the data, and then "learns" the underlying nonlinear transform...

...networks, each specializing in a particular domain of the feature space. The partitioning of the **data space** can be carried out using a clustering methodology. This local approach eliminates a significant portion...

2/3,K/17 (Item 5 from file: 654)  
 DIALOG(R)File 654:US Pat.Full.  
 (c) Format only 2004 The Dialog Corp. All rts. reserv.

4881383 \*\*IMAGE Available  
 Derwent Accession: 1998-287200

#### Utility

#### CERTIFICATE OF CORRECTION

E/ Method, system and computer program product for non-linear mapping of multi-dimensional data

Inventor: Agrafiotis, Dimitris K., Downingtown, PA  
 Lobanov, Victor S., Yardley, PA  
 Salemme, Francis R., Yardley, PA

Assignee: 3-Dimensional Pharmaceuticals, Inc. (02), Exton, PA  
 (Code: 39875)

Examiner: Follansbee, John (Art Unit: 211)

Assistant Examiner: Hirl, Joseph P.

Law Firm: Sterne, Kessler, Goldstein & Fox

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 6571227	A	20030527	US 99303671	19990503
CIP	US 6453246	A		US 9873845	19980507
CIP	US 6295514	A		US 97963872	19971111

Fulltext Word Count: 12397

#### Description of the Invention:

...alone for their conceptual elegance and ability to reproduce the topology and structure of the **data space** in a faithful and unbiased manner. Unfortunately, all known algorithms exhibit quadratic time complexity which...

...principle of probability sampling, the method employs an algorithm to multi-dimensionally scale a small **random sample**, and then "learns" the underlying non-linear transform using a multi-layer perceptron trained with...

2/3,K/18 (Item 6 from file: 654)  
 DIALOG(R)File 654:US Pat.Full.  
 (c) Format only 2004 The Dialog Corp. All rts. reserv.

4551847 \*\*IMAGE Available  
 Derwent Accession: 1999-395478

#### Utility



E/ **Weightless binary N-tuple thresholding hierarchies**

Inventor: King, Douglas B. S., Nr Preston, GB

MacDiarmid, Ian P., Preston, GB

Moore, Colin, Preston, GB

Assignee: BAE Systems plc (03), Farnborough, GB

BAE Systems PLC GB

Examiner: Malzahn, David H. (Art Unit: 211)

Law Firm: Nixon & Vanderhye P.C.

	Publication Number	Kind	Date	Application Number	Filing Date
	-----	--	-----	-----	-----
Main Patent	US 6272511	A	20010807	US 99366568	19990804
Continuation	Pending			WO 98GE3837	19981218
Priority				GB 9726752	19971219
				GB 9823382	19981027

Fulltext Word Count: 3279

Description of the Invention:

...of the first layer sum and threshold devices 14 of FIG. 2 is seen to  
**randomly sample the data space .**